

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:29 ; Search time 12.89 Seconds  
(without alignments)  
115,590 Million cell updates/sec

Title: US-09-829-481-4

Perfect score: 339  
Sequence: 1 MKSIAIIFVLVAFCLDGL.....IRRRGFCGRTFTTCVCYR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	20.2	72	4	US-09-003-198A-17
2	66	19.5	40	1	US-08-385-375-16
3	66	19.5	40	1	US-08-385-375-39
4	64.5	19.0	80	4	US-09-103-489-20
5	64	18.9	74	4	US-09-442-631-4
6	63.5	18.7	768	1	US-08-454-455-4
7	63.5	18.7	769	1	US-08-454-455-6
8	61.5	18.1	80	1	US-08-377-687-49
9	61.5	18.1	80	2	US-08-777-192-49
10	61.5	18.1	80	4	US-08-971-982-49
11	60.5	17.8	81	4	US-09-053-021-4
12	60.5	17.8	81	4	US-09-053-021-9
13	60.5	17.8	652	2	US-08-751-305-2
14	60	17.7	79	1	US-08-627-706-15
15	60	17.7	79	4	US-09-103-489-15
16	59.5	17.6	42	2	US-08-751-305-3
17	59.5	17.6	80	1	US-08-377-687-59
18	59.5	17.6	80	2	US-08-777-192-59
19	59.5	17.6	80	4	US-08-971-982-59
20	59	17.4	204	1	US-08-652-859-2
21	59	17.4	204	2	US-08-919-706-2
22	59	17.4	204	2	US-09-153-751-2
23	58.5	17.3	75	1	US-08-289-458-2
24	58.5	17.3	75	2	US-08-761-549-2
25	58.5	17.3	75	4	US-09-127-646-2
26	57.5	17.0	221	4	US-08-925-708-1
27	57	16.8	86	1	US-08-149-839B-14

28	57	16.8	86	1	US-08-451-568-14	Sequence 14, Appl
29	57	16.8	86	1	US-08-451-566-14	Sequence 14, Appl
30	57	16.8	86	2	US-08-777-113-14	Sequence 14, Appl
31	57	16.8	1394	6	5177197-30	Patent No. 5177197
32	57	16.8	1400	4	US-08-630-915A-37	Sequence 37, Appl
33	56.5	16.7	80	3	US-08-952-383A-14	Sequence 14, Appl
34	56.5	16.7	1248	4	US-08-882-046-6	Sequence 6, Appl
35	56	16.5	63	1	US-08-117-080-12	Sequence 12, Appl
36	56	16.5	63	1	US-08-471-329-12	Sequence 12, Appl
37	56	16.5	63	2	US-08-915-142-12	Sequence 12, Appl
38	56	16.5	67	4	US-09-053-021-2	Sequence 2, Appl
39	56	16.5	89	1	US-08-181-556-2	Sequence 2, Appl
40	56	16.5	103	1	US-08-486-013-20	Sequence 20, Appl
41	56	16.5	103	2	US-08-482-279-20	Sequence 20, Appl
42	56	16.5	103	2	US-08-342-268-20	Sequence 20, Appl
43	56	16.5	103	3	US-09-015-968-20	Sequence 20, Appl
44	56	16.5	103	4	US-09-397-386-20	Sequence 20, Appl
45	55.5	16.4	47	1	US-08-656-318A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1  
US-09-003-198A-17  
; Sequence 17, Application US/09003198A  
; Patent No. 6316407  
; GENERAL INFORMATION:  
; APPLICANT: Liang, Jihong  
; APPLICANT: Shah, Dilip Maganlal  
; APPLICANT: Wu, Yonnie S.  
; APPLICANT: Rosenberger, Cindy A.  
; APPLICANT: Hakimi, Salim  
; TITLE OF INVENTION: Antifungal Polypeptide and Methods for  
; TITLE OF INVENTION: Controlling Plant Pathogenic Fungi  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/003.198A  
; FILING DATE: 07-JAN-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Patterson, Melinda L.  
; REGISTRATION NUMBER: 33,062  
; REFERENCE/DOCKET NUMBER: MOBT:193  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (713) 787-1400  
; TELEFAX: (713) 787-1440  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 72 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-09-003-198A-17

Query Match 20.2%; Score 68.5; DB 4; Length 72;  
Best Local Similarity 28.4%; Pred No. 0.26;  
Matches 19; Conservative 11; Mismatches 24; Indels 13; Gaps 3;  
QY 2 KSIAIIFVLVAFCLDGLVEAGFCGPFNACK-----CHRHCKSIRRRG---GFCR 50

Db 4 KSLACLSELLLVLFAQELTVSEANTCENLAGYKGVCFGGCDRHCRRT--QEGAISSGRCR 61  
QY 51 GTFRITC 57  
Db 62 DDFRCWC 68

RESULT 2  
US-08-385-375-16  
; Sequence 16, Application US/08385375  
; Patent No. 5631144  
; GENERAL INFORMATION:  
; APPLICANT: LEMOINE, Yves  
; APPLICANT: NGUYEN, Martine  
; APPLICANT: ACHSTETTER, Tilman  
; APPLICANT: REICHART, Jean-Marc  
; TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A  
; TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF  
; TITLE OF INVENTION: MATURE PROTEINS BY RECOMINANT YEASTS, EXPRESSION  
; TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS FOR  
; TITLE OF INVENTION: THE PREPARATION OF PROTEINS  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/385.375  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/178,356  
; FILING DATE: 04-JAN-1994  
; APPLICATION NUMBER: FR 89/05687  
; FILING DATE: 28-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/FR90/00306  
; FILING DATE: 27-APR-1990  
; NAME: Crane-Feury, Sharon E  
; REGISTRATION NUMBER: 36,113  
; REFERENCE/DOCKET NUMBER: 017753-009  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-385-375-16

Query Match 19.5%; Score 66; DB 1; Length 40;  
Best Local Similarity 39.4%; Pred. No. 0.28;  
Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;  
QY 27 GCPFNAGKCHRRCKHSIRRRGGFCRGCTFRITTCVC 59  
Db 8 GTGINHSACAAHCLLRGNRGVCNG--KGVCVC 38

RESULT 3  
US-08-385-375-39  
; Sequence 39, Application US/08385375  
; Patent No. 5631144  
; GENERAL INFORMATION:  
; APPLICANT: LEMOINE, Yves  
; APPLICANT: NGUYEN, Martine  
; APPLICANT: ACHSTETTER, Tilman  
; APPLICANT: REICHART, Jean-Marc  
; TITLE OF INVENTION: APPLICATION OF NOVEL DNA FRAGMENTS AS A  
; TITLE OF INVENTION: SEQUENCE CODING FOR A SIGNAL PEPTIDE FOR THE SECRETION OF  
; TITLE OF INVENTION: MATURE PROTEINS BY RECOMINANT YEASTS, EXPRESSION  
; TITLE OF INVENTION: CASSETTES, TRANSFORMED YEASTS AND CORRESPONDING PROCESS FOR  
; TITLE OF INVENTION: THE PREPARATION OF PROTEINS  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/385.375  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/178,356  
; FILING DATE: 04-JAN-1994  
; APPLICATION NUMBER: FR 89/05687  
; FILING DATE: 28-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/FR90/00306  
; FILING DATE: 27-APR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crane-Feury, Sharon E  
; REGISTRATION NUMBER: 36,113  
; REFERENCE/DOCKET NUMBER: 017753-009  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 40 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-385-375-39

Query Match 19.5%; Score 66; DB 1; Length 40;  
Best Local Similarity 39.4%; Pred. No. 0.28;  
Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;  
QY 27 GCPFNAGKCHRRCKHSIRRRGGFCRGCTFRITTCVC 59  
Db 8 GTGINHSACAAHCLLRGNRGVCNG--KGVCVC 38

RESULT 4  
US-09-103-489-20  
; Sequence 20, Application US/09103489  
; Patent No. 6215048  
; GENERAL INFORMATION:  
; APPLICANT: Liang, Jihongq  
; APPLICANT: Shah, Dilip M.  
; APPLICANT: Wu, Yonnie S.

```

RESULT      5
US-09-442-631-4
; Sequence 4, Application US/09442631
; Patent No. 6300489
; GENERAL INFORMATION:
; APPLICANT: OH, BOUNG-JUN
; APPLICANT: KO, MOON KYUNG
; APPLICANT: SHIN, BYONGCHUL
; APPLICANT: CHUNG, CHANG HO
; TITLE OF INVENTION: SMALL AND CYSTEINE RICH ANTIFUNGAL DEFENSIN AND
; TITLE OF INVENTION: THIONIN-LIKE PROTEIN GENES HIGHLY EXPRESSED IN THE
; TITLE OF INVENTION: INCOMPATIBLE INTERACTION
; FILE REFERENCE: 1942/44
; CURRENT APPLICATION NUMBER: US/09/442,631
; CURRENT FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Capsicum annuum
US-09-442-631-4

```

```

Query Match      18.78; Score 63.5; DB 1; Length 768;
Best Local Similarity 24.78; Pred. No. 10;
Matches 19; Conservative 8; Mismatches 11; Indels 39; Gaps
5;

Qy 14 PCILEDGIVEAGFGCPFN-----AGKCH-----RCKKSTR 44
      | | | | | | | | | | | | | | | | | | | | | |
Db 545 YCEKID-----FSCPYHHSLCAGHGECEAGRCQCFSGWGEDRCQCPSAQAHC--VNS 596

Qy 45 RGGFCRGFTRTTCVYR 61

```



```
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777,192
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-777-192-49

Query Match 18.1%; Score 61.5; DB 2; Length 80;
Best Local Similarity 30.0%; Pred. No. 1.9;
Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 4;

QY 4 IAIIFVLVAFCLILE-DGIVEAGFGCPFNAG-----KCHRHCKSIIR-RGGFCRG 51
||:| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 IALLFAALVLEAFAPTMTVEAQKLCERPSTGWSGCGNNACKNQCNLEKARHGSNY 67

-QY 52 TFRT-TCVCY 60
| | | | |
Db 68 VFPAHKCICY 77

RESULT 10
US-08-971-982-49
; Sequence 49, Application US/08971982
; Patent No. 6187904
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; CAMMUE, BRUNO P.A.
; OSBORN, RUPERT W.
; REES, SARAH B.
; TERRAS, FRANKY R.G.
; VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,982
; FILING DATE: 17-NOV-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-08-971-982-49

Query Match 18.1%; Score 61.5; DB 4; Length 80;
Best Local Similarity 30.0%; Pred. No. 1.9;
Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 4;

QY 4 IAIIFVLVAFCLILE-DGIVEAGFGCPFNAG-----KCHRHCKSIIR-RGGFCRG 51
||:| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 IALLFAALVLEAFAPTMTVEAQKLCERPSTGWSGCGNNACKNQCNLEKARHGSNY 67

QY 52 TFRT-TCVCY 60
| | | | |
Db 68 VFPAHKCICY 77

RESULT 11
US-09-053-021-4
; Sequence 4, Application US/09053021
; Patent No. 6270785
; GENERAL INFORMATION:
; APPLICANT: SELISKO, Barbara
; APPLICANT: GARCIA-RODRIGUEZ, Consuelo
; APPLICANT: ZAMUDIO-ZUNIGA, Fernando
; APPLICANT: BECERRIL-LUJAN, Baltazar
; APPLICANT: POSSANI-POSTAY, Lourival D.
; TITLE OF INVENTION: Primary Sequence and cDNA of
; Patent No. 6270785
; TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the
; Genus Centruroides
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,021
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,007
; FILING DATE: 30-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/848,261
; FILING DATE: 29-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett, Arthur S.
; REGISTRATION NUMBER: 20,338
```

; REFERENCE/DOCKET NUMBER: 06899.0001-01000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)408-4000  
; TELEFAX: (202)408-4400  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 81 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-053-021-4

Query Match 17.8%; Score 60.5; DB 4; Length 81;  
Best Local Similarity 23.4%; Pred. No. 2.4;  
Matches 15; Conservative 10; Mismatches 28; Indels 11; Gaps 3;  
QY 4 IAIIFIVLAFCLIEDGIVEAGFGCPFNA-----GKCHRHCKSIIRRG--GFCRGTFRT 55  
Db 1 ITACLVLTGTVCAKEGYLVNKSTGCKYNCLILGENKNCDMECKAKNOGGSYGYC---YKL 57  
QY 56 TCVC 59  
Db 58 ACWC 61

RESULT 12  
US-09-053-021-9  
; Sequence 9, Application US/09053021  
; Patent No. 6270785  
; GENERAL INFORMATION:  
; APPLICANT: SELISKO, Barbara  
; APPLICANT: GARCIA-RODRIGUEZ, Consuelo  
; APPLICANT: ZAMUDIO-ZUNIGA, Fernando  
; APPLICANT: BECERRIL-LUJAN, Baltazar  
; APPLICANT: POSSANI-POSTAV, Lourival D.  
; TITLE OF INVENTION: Primary Sequence and cDNA of  
; Patent No. 6270785  
; TITLE OF INVENTION: Insecticidally Effective Toxins from Scorpions of the  
; TITLE OF INVENTION: Genus Centruroides  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSEE: Dunner, L.L.P.  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/053,021  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/017,007  
; FILING DATE: 30-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/848,261  
; FILING DATE: 29-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Garrett, Arthur S.  
; REGISTRATION NUMBER: 20,338  
; REFERENCE/DOCKET NUMBER: 06899.0001-01000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)408-4000  
; TELEFAX: (202)408-4400  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 81 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-053-021-9

Query Match 17.8%; Score 60.5; DB 4; Length 81;  
Best Local Similarity 23.4%; Pred. No. 2.4;  
Matches 15; Conservative 10; Mismatches 28; Indels 11; Gaps 3;  
QY 4 IAIIFIVLAFCLIEDGIVEAGFGCPFNA-----GKCHRHCKSIIRRG--GFCRGTFRT 55  
Db 1 ITACLVLTGTVCAKEGYLVNKSTGCKYNCLILGENKNCDMECKAKNOGGSYGYC---YKL 57  
QY 56 TCVC 59  
Db 58 ACWC 61

RESULT 13  
US-08-751-305-2  
; Sequence 2, Application US/08751305  
; Patent No. 5965439  
; GENERAL INFORMATION:  
; APPLICANT: Tenner et al., Andrea J.  
; TITLE OF INVENTION: HOST DEFENSE ENHANCEMENT  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/751,305  
; FILING DATE: 18-NOV-1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell, Jr., John R.  
; REGISTRATION NUMBER: 31,678  
; REFERENCE/DOCKET NUMBER: 07306/012001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 652 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-751-305-2

Query Match 17.8%; Score 60.5; DB 2; Length 652;  
Best Local Similarity 36.2%; Pred. No. 19;  
Matches 17; Conservative 2; Mismatches 13; Indels 15; Gaps 3;  
QY 22 VFAGFGCPFNAGKCHRHCKSIIRRG-----FCRGTFR-----TTC 57  
Db 258 VSPKYGCNENGGCHQDC-----FEGDGSFLCGCRPGFRLLDLVTC 300  
RESULT 14  
US-08-627-706-15  
; Sequence 15, Application US/08627706  
; Patent No. 5773696

```
; GENERAL INFORMATION:
; APPLICANT: Liang, Jihong
; APPLICANT: Shah, Dilip M.
; APPLICANT: Wu, Yonnie S.
; APPLICANT: Rosenberger, Cindy A.
; TITLE OF INVENTION: Antifungal Polypeptide and Methods for
; TITLE OF INVENTION: Controlling Plant Pathogenic Fungi
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F
; STREET: 700 Chesterfield Village Parkway No. 5773696th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,706
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Cohen, Charles E.
; REGISTRATION NUMBER: 34,565
; REFERENCE/DOCKET NUMBER: 38-21(10700)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6224
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-627-706-15
```

```
Query Match 17.7% Score 60; DB 1; Length 79;
Best Local Similarity 27.5%; Pred. No. 2.7;
Matches 19; Conservative 9; Mismatches 29; Indels 12; Gaps 4;

QY 4 IAIIFIVLVAFCILE-DGIVEAGF-----GCPFNAGKCHRHCKSIRR-RGGFCRGT 52
|:::| | | | | | | | | | | | | | | | | | | | |
Db 8 ISLLFAALVLFAPFAEPTMVDARLCERPSTGWSGCGNNNACRNLERAHSGSNV 67

QY 53 FRT-TCVCY 60
| | | | |
Db 68 FPAHKCICY 76

RESULT 15
US-09-103-489-15
; Sequence 15, Application US/09103489
; Patent No. 6215048
; GENERAL INFORMATION:
; APPLICANT: Liang, Jihong
; APPLICANT: Shah, Dilip M.
; APPLICANT: Wu, Yonnie S.
; APPLICANT: Rosenberger, Cindy A.
; TITLE OF INVENTION: Antifungal Polypeptide and Methods for
; TITLE OF INVENTION: Controlling Plant Pathogenic Fungi
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles E. Cohen, Monsanto Company, BB4F
; STREET: 700 Chesterfield Village Parkway No. 6215048th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/103,489
; FILING DATE: 24-JUN-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Cohen, Charles E.
; REGISTRATION NUMBER: 34,565
; REFERENCE/DOCKET NUMBER: 38-21 (10700)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 537-6224
; TELEFAX: (314) 537-6047
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-103-489-15

Query Match 17.7% Score 60; DB 4; Length 79;
Best Local Similarity 27.5%; Pred. No. 2.7;
Matches 19; Conservative 9; Mismatches 29; Indels 12; Gaps 4;

QY 4 IAIIFIVLVAFCILE-DGIVEAGF-----GCPFNAGKCHRHCKSIRR-RGGFCRGT 52
|:::| | | | | | | | | | | | | | | | | | | | |
Db 8 ISLLFAALVLFAPFAEPTMVDARLCERPSTGWSGCGNNNACRNLERAHSGSNV 67

QY 53 FRT-TCVCY 60
| | | | |
Db 68 FPAHKCICY 76

Search completed: September 17, 2002, 15:44:35
Job time: 126 sec
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:09 ; Search time 29.91 seconds  
(without alignments)  
226.530 Million cell updates/sec

Title: US-09-829-481-4  
Perfect score: 339  
Sequence: 1 MKSAIIIFVLVAFCLLEDG.....IRRRGGFCRGFTTTCVCYR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*

```

1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	177	52.2	38	19 AAW66437	Cationic peptide s
2	177	52.2	38	21 AAY91736	Cationic peptide s
3	133.5	39.4	38	15 AAR51160	Antibacterial pept
4	133.5	39.4	38	19 AAW66436	Cationic peptide i
5	133.5	39.4	38	21 AAY91735	Cationic peptide i
6	71.5	21.1	43	16 AAR82841	Allomycin A antiba
7	70	20.6	75	22 AAE10361	Soybean Gly m2 pro
8	70	20.6	94	10 AAP94261	Antibacterial poly
9	69.5	20.5	43	16 AAR82842	Allomycin B antiba
10	68.5	20.2	40	14 AAR30562	Sapecin - an antib
11	68.5	20.2	72	21 AAY44509	Pea Defensin prote

12	68	20.1	40	19 AAW66435	Cationic peptide s
13	68	20.1	40	21 AAY91734	Cationic peptide s
14	68	20.1	77	21 AAB24718	Plant SDF encoded
15	68	20.1	78	21 AAB24718	Plant SDF encoded
16	68	20.1	94	10 AAP90001	Antibacterial poly
17	68	20.1	100	21 AAB24717	Plant SDF encoded
18	67.5	19.9	43	18 AAW26000	Tenebrion. Tenebrio
19	67	19.8	567	21 AAB24967	Plant SDF encoded
20	67	19.8	799	21 AAB24966	Plant SDF encoded
21	67	19.8	817	21 AAB24965	Plant SDF encoded
22	66	19.5	40	11 AAR00720	Phormia terranova
23	66	19.5	40	19 AAW66454	Cationic peptide p
24	66	19.5	40	21 AAY91753	Cationic peptide p
25	66	19.5	45	15 AAR57129	Yeast pro-AMF C-te
26	65.5	19.3	84	21 AAG15344	Arabidopsis thalia
27	64.5	19.0	80	18 AAW19280	Raphanus sativus a
28	64.5	19.0	80	18 AAW19617	Radish antifungal
29	64.5	19.0	85	22 AAB20070	Scorpion sodium ch
30	63.5	18.7	54	22 AAU52922	Propionibacterium
31	63.5	18.7	80	22 AAB61792	Radish antifungal
32	63.5	18.7	81	20 AAY28848	Maize basal endosp
33	63.5	18.7	768	13 AAR27683	Rabbit beta-8 inte
34	63.5	18.7	768	13 AAR27684	Human beta-8 inte
35	63	18.6	88	21 AAG10630	Arabidopsis thalia
36	63	18.6	88	21 AAG45781	Arabidopsis thalia
37	63	18.6	90	21 AAG10629	Arabidopsis thalia
38	63	18.6	90	21 AAG45780	Arabidopsis thalia
39	63	18.6	106	21 AAG10628	Arabidopsis thalia
40	62	18.3	106	21 AAG45779	Arabidopsis thalia
41	62	18.3	27	21 AAY88036	Human growth hormo
42	61.5	18.1	77	21 AAG04820	Arabidopsis thalia
43	61.5	18.1	77	21 AAG38373	Arabidopsis thalia
44	61.5	18.1	77	21 AAY68682	Amino acid sequenc
45	61.5	18.1	80	14 AAR33706	Antifungal protein

#### ALIGNMENTS

RESULT 1  
AAW66437  
ID AAW66437 standard; peptide: 38 AA.  
AC AAW66437;  
XX  
XX  
DT 12-JAN-1999 (first entry)  
XX  
XX Cationic peptide scorpion defensin.  
DE  
XX

Indolicidin analogue; resistance; cationic peptide; antibiotic;  
bacterial infection; tolerance; antibacterial; microorganism;  
bacteria; fungus; parasite; virus.

OS Lelurus quinquestriatus.

PN WO9840401-A2.

PD 17-SEP-1998.

PF 10-MAR-1998; 98WO-CA00190.

PR 25-FEB-1998; 98US-0030619.

PR 10-MAR-1997; 97US-0040649.

PR 20-AUG-1997; 97US-0915314.

PR 26-SEP-1997; 97US-0060099.

PA (MTCR-) MICROLOGIX BIOTECH INC.

PI Fraser JR, McNicol PJ, West MHP;

XX WPI; 1998-520800/44.

XX New Indolicidin peptide analogues - useful for, e.g. enhancing

PT activity of antibiotic or overcoming tolerance, acquired resistance  
 XX or inherent resistance of microorganisms  
 PS Disclosure; Page 10; 105pp; English.

XX AAW66393 to AAW66469 represent native cationic peptides from the  
 CC present invention. The present invention describes compositions and  
 CC methods for treating infection, especially bacterial infections. The  
 CC compositions and methods use cationic peptides in combination with an  
 CC antibiotic agent which are then administered to a patient to enhance the  
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)  
 CC acquired resistance; and (c) inherent resistance. The combinations of  
 CC antibiotics and cationic peptides can provide synergistic activity  
 CC against a microorganism that is tolerant, inherently resistant, or has  
 CC acquired resistance to an antibiotic agent. They can be used for killing  
 CC e.g. bacteria, fungi, parasites and viruses.

XX Sequence 38 AA;

Query Match 52.2%; Score 177; DB 19; Length 38;  
 Best Local Similarity 73.0%; Pred. No. 1.9e-13;  
 Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 25 GFGCPFNAGKCHRHCKSIIRRGFGCRGTFRTTCVCYR 61  
 ||||| | |||||:|||||: | | | | |  
 Db 1 gfgcplnggachrhersirrrggycagffkqtctcyr 37

#### RESULT 2

AAAY91736  
 ID AAY91736 standard; Peptide; 38 AA.

XX AAY91736;

XX 06-JUN-2000 (first entry)

DE Cationic peptide scorpion defensin amino acid sequence.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;  
 KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;  
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;  
 KW multidrug resistance.

XX Unidentified.

PN W09965506-A2.

XX 23-DEC-1999.

PF 14-JUN-1999; 99WO-CA00552.

PR 12-JUN-1998; 98US-0096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

XX WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated  
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumours

XX Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which  
 CC can be used in the pharmaceutical composition of the invention. The  
 CC invention relates to a pharmaceutical composition containing at least one  
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The  
 CC modification of peptides with APO increases their activity against tumour  
 CC cells, including those with a multidrug resistant phenotype. The  
 CC pharmaceutical composition can be used to treat tumours, specifically

CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
 CC cervix, uterus, skin, prostate, liver and colon.

SQ Sequence 38 AA;

Query Match 52.2%; Score 177; DB 21; Length 38;  
 Best Local Similarity 73.0%; Pred. No. 1.9e-13;  
 Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 25 GFGCPFNAGKCHRHCKSIIRRGFGCRGTFRTTCVCYR 61  
 ||||| | |||||:|||||: | | | | |  
 Db 1 gfgcplnggachrhersirrrggycagffkqtctcyr 37

#### RESULT 3

AAR51160  
 ID AAR51160 standard; peptide; 38 AA.

XX AAR51160;

XX 18-OCT-1994 (first entry)

XX Antibacterial peptide induced in *Aeschna cyanea*.

DE Defensin; antibacterial peptide; Odonata; paleopteran insect;  
 KW dragonfly.

XX *Aeschna cyanea*.

XX Key Location/Qualifiers

FT Disulfide-bond 4..26

FT Disulfide-bond 11..34

FT Disulfide-bond 14..36

FT Region 11..21

FT /label= alpha\_helix

FT /note= "approximate position"

FT Region 24..38

FT /label= beta-pleated\_sheet

FT /note= "approximate position"

PN FR2695392-A.

PD 11-MAR-1994.

XX 04-SEP-1992; 92FR-0010609.

XX 04-SEP-1992; 92FR-0010609.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX Bulet P, Hetru C, Hoffmann J;

XX WPI; 1994-103535/13.

XX New antibacterial peptide(s) from dragonfly - for medical,  
 PT veterinary, agricultural and food preservation use

XX Claim 10; Page 17; 25pp; French.

XX A defensin is induced in the dragonfly *Aeschna cyanea* by bacterial  
 CC infection. The defensin (AAR51160) contains 3 disulphide bonds and has  
 CC a structure comprising an alpha helix linked to a C-terminal  
 CC antiparallel beta-pleated sheet through 2 of the disulphide bonds.  
 CC The peptide has antibacterial activity which makes it suitable for  
 CC use in the treatment of bacterial infections of eyes and ears and  
 CC for oral/dental hygiene and gynaecology.

XX Sequence 38 AA;

Query Match 39.4%; Score 133.5; DB 15; Length 38;  
 Best Local Similarity 55.3%; Pred. No. 1.9e-08;

Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPNAGKCHRHCKSIR-RRGGFCRGTFRTTCVCYR 61  
 ||||| : :|||||:| | ||| : | |||  
 Db 1 gfgcpIdmqchrcqtigrsggycsgplkltctcyr 38

## RESULT 4

AAW66436  
 ID AAW66436 standard; peptide; 38 AA.

XX AAW66436;

XX 12-JAN-1999 (first entry)

XX Cationic peptide insect defensin.

XX Indolicidin analogue; resistance; cationic peptide; antibiotic;

KW bacterial infection; tolerance; antibacterial; microorganism;

KW bacteria; fungus; parasite; virus.

XX Aeschna cyanae.

XX WO9840401-A2.

XX 17-SEP-1998.

XX 10-MAR-1998; 98WO-CA00190.

XX 25-FEB-1998; 98US-0030619.

XX 10-MAR-1997; 97US-0040649.

XX 20-AUG-1997; 97US-0915314.

XX 26-SEP-1997; 97US-0060099.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Fraser JR, McNicol PJ, West MHP;

XX WPI; 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing  
 PT activity of antibiotic or overcoming tolerance, acquired resistance  
 or inherent resistance of microorganisms

XX Disclosure; Page 10; 105pp; English.

XX AAW66393 to AAW66469 represent native cationic peptides from the  
 CC present invention. The present invention describes compositions and  
 CC methods for treating infection, especially bacterial infections. The  
 CC compositions and methods use cationic peptides in combination with an  
 CC antibiotic agent which are then administered to a patient to enhance the  
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)  
 CC acquired resistance; and (c) inherent resistance. The combinations of  
 CC antibiotics and cationic peptides can provide synergistic activity  
 CC against a microorganism that is tolerant, inherently resistant, or has  
 CC acquired resistance to an antibiotic agent. They can be used for killing  
 CC e.g. bacteria, fungi, parasites and viruses.

XX Sequence 38 AA;

Query Match 39.4%; Score 133.5; DB 19; Length 38;  
 Best Local Similarity 55.3%; Pred. No. 1.9e-08;  
 Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPNAGKCHRHCKSIR-RRGGFCRGTFRTTCVCYR 61  
 ||||| : :|||||:| | ||| : | |||  
 Db 1 gfgcpIdmqchrcqtigrsggycsgplkltctcyr 38

## RESULT 5

AAAY91735  
 ID AAY91735 standard; Peptide; 38 AA.

XX AAY91735;  
 AC 06-JUN-2000 (first entry)  
 DT Cationic peptide insect defensin amino acid sequence.  
 DE  
 DE  
 XX  
 XX  
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;  
 KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;  
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;  
 KW multidrug resistance.

XX Unidentified.

XX WO9965506-A2.

XX 23-DEC-1999.

XX 14-JUN-1999; 99WO-CA00552.

XX 12-JUN-1998; 98US-0096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

XX Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

XX WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated  
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumours  
 PT  
 XX Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which  
 CC can be used in the pharmaceutical composition of the invention. The  
 CC invention relates to a pharmaceutical composition containing at least one  
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The  
 CC modification of peptides with APO increases their activity against tumour  
 CC cells, including those with a multidrug resistant phenotype. The  
 CC pharmaceutical composition can be used to treat tumours, specifically  
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
 CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 38 AA;

Query Match 39.4%; Score 133.5; DB 21; Length 38;  
 Best Local Similarity 55.3%; Pred. No. 1.9e-08;  
 Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPNAGKCHRHCKSIR-RRGGFCRGTFRTTCVCYR 61  
 ||||| : :|||||:| | ||| : | |||  
 Db 1 gfgcpIdmqchrcqtigrsggycsgplkltctcyr 38

## RESULT 6

AAAR82841  
 ID AAR82841 standard; peptide; 43 AA.

XX AAR82841;

XX 25-MAR-1996 (first entry)

XX Allomycin A antibacterial peptide.

XX Allomycin; antibacterial; drug; food; recombinant.

XX Allomyrina dichotoma.

XX JP07196688-A.

XX 01-AUG-1995.

XX PF 28-DEC-1993; 93JP-0350294.  
 XX PR 28-DEC-1993; 93JP-0350294.  
 XX PA (AMAN ) AMANO PHARM KK.  
 XX DR WPI; 1995-299564/39.  
 XX DR N-PSDB; AAT01306.  
 PT Novel polypeptide derived from Allomyrina dichotoma larvae - is  
 PT physiologically active against bacteria, and may be used in drugs  
 PT and foodstuffs  
 XX Claim 1; Page 10; 14pp; Japanese.  
 XX Two new peptides have been derived from Allomyrina dichotoma,  
 CC allomycin A and allomycin B (AAR81841 and AAR82842). The new peptides  
 CC have antibacterial activity which makes them useful in the drug and  
 CC foodstuff industries. The proteins may be prepared using DNA coding  
 CC for the peptides (AAT01306 and AAT01307) in standard recombinant  
 CC techniques, this also enables the large scale production of large  
 CC quantities of the peptides.  
 XX Sequence 43 AA;  
 SQ  
 Query Match 21.1%; Score 71.5; DB 16; Length 43;  
 Best Local Similarity 42.9%; Pred. No. 0.28;  
 Matches 15; Conservative 2; Mismatches 15; Indels 3; Gaps 1;  
 QY 27 GCPFNAGKCHRRCKSIIRRGFCRGTFRFTTCVCYR 61  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 12 gfaahnsicaahclaigrsgsce---Igvclcrr 43  
 | | | | | | | | | | | | | | | | | | | | | |  
 RESULT 7  
 AAEL0361  
 ID AAEL0361 standard; Protein; 75 AA.  
 AC AAEL0361;  
 DT 10-DEC-2001 (first entry)  
 XX Soybean Gly m2 protein from clone s1slc.pk027.all.  
 DE Soybean; allergen; transgenic plant; P34 protein; Gly m Bd 30K; Gly m 1;  
 KW soybean vacuolar protein; Gly m 1A; Gly m 1B; rGLY m3; Glycinin G1;  
 KW alalb; food; infant formula; animal feed; coating; salad oil; syrup;  
 KW spraying oil; roasting oil; frying oil; cracker; confectionery product;  
 KW snack food; topping; sauce; batter; bread; bread; bread; bread; bread;  
 KW Gly m2 protein; clone s1slc.pk027.all.  
 XX Glycine max.  
 OS  
 XX WO200168887-A2.  
 PN 20-SEP-2001.  
 XX 15-MAR-2001; 2001WO-US08254.  
 PF 16-MAR-2000; 2000US-0189823.  
 PR (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX (PTON-) PIONEER HI-BREED INT INC.  
 PA Jung R, Kinney AJ;  
 PI WPI; 2001-582460/65.  
 XX DR N-PSDB; AADI7530.  
 DR Recombinant expression construct to lower allergen (e.g. Gly m Bd 30K)  
 PT content of a soybean, comprises a nucleic acid fragment encoding the

PT allergen, useful for producing soybean plants which can be used to make  
 PT soybean products  
 XX Claim 56; Page 52; 57pp; English.  
 XX The patent discloses hypoallergenic transgenic soybeans and recombinant  
 CC expression constructs to lower soybean vacuolar protein, commonly known  
 CC as P34 (Gly m Bd 30K or Gly m 1) and other allergens such as Gly m 1A,  
 CC Gly m 1B, rGLY m3 or Glycinin G1 (alalb). The allergen content of the  
 CC soybean is reduced by sense suppression which is accomplished by using  
 CC the expression construct that comprises a nucleic acid fragment encoding  
 CC the allergen. The constructs are useful for producing hypoallergenic  
 CC transgenic soybean plants which can be used to make hypoallergenic  
 CC soybean products which can be used in a variety of food (e.g. infant  
 CC formulas) and animal feed applications. The oil made from seeds of the  
 CC hypoallergenic transgenic soybean plants can be used as ingredients,  
 CC as coatings, as salad oils, as spraying oils, as roasting oils, and  
 CC as frying oils. The foods in which the oil may be used include crackers  
 CC and snack foods, confectionery products, syrups and toppings, sauces,  
 CC batter and bread mixtures, baking mixes and doughs. The present  
 CC sequence is a soybean Gly m2 protein from s1slc.pk027.all clone. This  
 CC sequence is a minor soybean seed allergen.  
 XX Sequence 75 AA;  
 SQ  
 Query Match 20.6%; Score 70; DB 22; Length 75;  
 Best Local Similarity 32.4%; Pred. No. 0.73;  
 Matches 23; Conservative 7; Mismatches 23; Indels 18; Gaps 4;  
 QY 2 KSIA-IIIFVLVAFCLIE-----DGIVEAGFCPFNAGKCHRRCKS----IRRRG 46  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 4 ksiaqlcflflvlfaqvqvteaktcoenladtyrgpcttgscdhckckehlir---- 60  
 | | | | | | | | | | | | | | | | | | | | | |  
 QY 47 GFCRGTRFTTC 57  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 61 gcrddfrwc 71  
 | | | | | | | | | | | | | | | | | | | | | |  
 RESULT 8  
 AAP94261  
 ID AAP94261 standard; protein; 94 AA.  
 XX AAP94261;  
 AC AAP94261;  
 DT 17-JUN-1990 (first entry)  
 XX Antibacterial polypeptide precursor of Sarcophaga peregrina.  
 DE Antibacterial polypeptide precursor of Sarcophaga peregrina.  
 KW Antibacterial polypeptide precursor; Sarcophaga peregrina.  
 XX Sarcophaga peregrina.  
 OS  
 XX Key Location/Qualifiers  
 FH Region 55..94  
 FT /note="cloned antibacterial polypeptide"  
 FT Region 24..62  
 FT /note="previously disclosed sequence"  
 FT Region 1..23  
 FT /label=signal sequence  
 XX EP303858-A.  
 PN 22-FEB-1989.  
 XX 26-JUL-1988; 88EP-0112072.  
 PF 20-AUG-1987; 87JP-0205165.  
 PR (SANW ) SANWA KAGAKU KENKYUSHO CO. LTD.  
 XX Sawai K, Natori S, Takahashi H, Tanaka K, Mitani T, Kurono M;  
 PI



```
XX 13-JAN-2000.
PD XX
PF PF
XX 02-JUL-1999; 99WO-CA00608.
XX
PR 03-JUL-1998; 98CA-2242116.
PR 06-JUL-1998; 98US-0091751.
XX
PA (UYMA-) UNIV MANITOBA.
XX
XX Fristensky B, Wang Y;
XX
DR WPI; 2000-126938/11.
DR N-PSDB; AA229412.
XX
XX Recombinant expression system for expressing DRR206 or defensin, used
PT to produce pathogen resistant Brassica napus -
PT
XX Claim 11; Fig 9; 39pp; English.
XX
CC The present amino acid sequence is the defensin protein, derived from
CC the clone Drr230 of pea. This sequence is induced by bacterial and
CC fungal pathogens like, blackleg fungus. The gene sequence is used in a
CC recombinant expression system, capable of transforming plants like,
CC Canola (Brassica napus), under the constitutive control of CaMV 35S
CC inducible promoter, responsive to pathogen infections. T-DNA sequence is
CC also present, for integration of the expression system into the plant
CC genome. Transgenic plants expressing DRR206 and defensin protein,
CC inhibited fungal growth in-vitro and are resistant to pathogenic
CC infections of Rhizoctonia solani, Leptosphaeria maculans and Sclerotinia
CC sclerotiorum.
XX
SQ Sequence 72 AA;

Query Match 20.2%; Score 68.5; DB 21; Length 72;
Best Local Similarity 28.4%; Pred. No. 1;
Matches 19; Conservative 11; Mismatches 24; Indels 13; Gaps 3;

QY 2 KSTAITIVLVPCLIEDGIVEAGFCPPNAGK-----CHRHCKSIIRRG---GFCR 50
DB ||:| :|: :|: || | ||:| :|: ||
4 kslacIsflllvfvaqeiVvseantcenlagsgkvcfggcdrhrt--qegalsgrcr 61

QY 51 GPFRTTC 57
DB || |
62 ddfrcwc 68

RESULT 12
AAW66435
ID AAW66435 standard; peptide; 40 AA.
XX
AC AAW66435;
XX
XX 12-JAN-1999 (first entry)
XX
DE Cationic peptide sapecin.
XX
XX Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW bacterial infection; tolerance; antibacterial; microorganism;
KW bacteria; fungus; parasite; virus.
XX
OS Sacrophaga peregrina.
XX
PN WO9840401-A2.
XX
PD 17-SEP-1998.
XX
PF 10-MAR-1998; 98WO-CA00190.
XX
XX 25-FEB-1998; 98US-0030619.
PR 10-MAR-1997; 97US-0040649.
PR 20-AUG-1997; 97US-0915314.
XX

26-SEP-1997; 97US-0060099.
(MICR-) MICROLOGIX BIOTECH INC.
Fraser JR, McNicol PJ, West MHP;
WPI; 1998-520800/44.
New indolicidin peptide analogues - useful for, e.g. enhancing
activity of antibiotic or overcoming tolerance, acquired resistance
or inherent resistance of microorganisms
Disclosure; Page 10; 105pp; English.
AAW66393 to AAW66469 represent native cationic peptides from the
present invention. The present invention describes compositions and
methods for treating infection, especially bacterial infections. The
compositions and methods use cationic peptides in combination with an
antibiotic agent which are then administered to a patient to enhance the
activity of the antibiotic agent, to overcome: (a) tolerance; (b)
acquired resistance; and (c) inherent resistance. The combinations of
antibiotics and cationic peptides can provide synergistic activity
against a microorganism that is tolerant, inherently resistant, or has
acquired resistance to an antibiotic agent. They can be used for killing
e.g. bacteria, fungi, parasites and viruses.
SQ Sequence 40 AA;

Query Match 20.1%; Score 68; DB 19; Length 40;
Best Local Similarity 39.4%; Pred. No. 0.65; Mismatches 16; Indels 2; Gaps 1;
Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

QY 27 GCFPNAGKCHRHCKSIIRRGGFCGRTTCVC 59
DB | | | | | | | | | | | | | | | |
8 gTgInhsacaahcHllrgnrggycng--kavcvc 38

RESULT 13
AAW91734
ID AAW91734 standard; Peptide; 40 AA.
XX
AC AAW91734;
XX
XX 06-JUN-2000 (first entry)
XX
DE Cationic peptide Sapecin amino acid sequence.
XX
KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma,
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.
XX
OS Unidentified.
XX
PN WO9965506-A2.
XX
PD 23-DEC-1999.
XX
PF 14-JUN-1999; 99WO-CA00552.
XX
PR 12-JUN-1998; 98US-0096541.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
WPI; 2000-223549/19.
XX
XX Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
PT
XX
```

PS Disclosure; Page 11; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which  
CC can be used in the pharmaceutical composition of the invention. The  
CC invention relates to a pharmaceutical composition containing at least one  
CC activated polyoxyalkylene (APO)-modified cationic peptide. The  
CC modification of peptides with APO increases their activity against tumour  
CC cells, including those with a multidrug resistant phenotype. The  
CC pharmaceutical composition can be used to treat tumours, specifically  
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 40 AA;

Query Match 20.1%; Score 68; DB 21; Length 40;

Best Local Similarity 39.4%; Pred. No. 0.65; Length 40;

Matches 13; Conservative 2; Mismatches 16; Indels 2; Gaps 1;

Oy 27 GCFNAGKCHRHCKSIRRRGGFCRGTFRTTCVC 59

Db 8 gtginhsacaahcillrgrngycng--kavcvc 38

RESULT 14

AAB24719

ID AAB24719 standard; Peptide; 77 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

CC polynucleotide sequences and polypeptides encoded by them given in the  
CC present invention.

SQ Sequence 77 AA;

Query Match 20.1%; Score 68; DB 21; Length 77;

Best Local Similarity 30.1%; Pred. No. 1.3; Length 77;

Matches 22; Conservative 9; Mismatches 26; Indels 16; Gaps 4;

Oy 2 KSTAIIFIVLVAFCILEDGIV-EAGFGCCPFN-----AGKCHRHCKSIRRRG-G 47

Db 3 kllqvstvmiiftlilvgvvaneglgpkkqcnellkqsnvcvaecdsmcvkkrkgag 62

Oy 48 FCRGTFRTTCVCY 60

Db 63 ycspsk--cycy 73

RESULT 15

AAB24718

ID AAB24718 standard; Peptide; 78 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;

PI Zheng L;

PI WPI; 2000-465970/40.

PI New corn plant and Arabidopsis thaliana sequence-determined DNA

PI fragments, useful for expressing gene products and for controlling

PI expression of a target gene -

PI Claim 14; Page 400; 673pp; English.

PI The present invention describes polynucleotides, such as complete cDNA

PI sequences and/or sequences of genomic DNA encompassing complete genes,

PI portions of genes, and/or intergenic regions, collectively referred to

PI as sequence-determined DNA fragments (SDFs), from corn plants and

PI Arabidopsis thaliana. The SDFs are promoters, structural genes,

PI untranslated regions (UTRs), or 3' termination sequences. They can be

PI used for expressing a gene product and controlling expression of a

PI target gene, either as a promoter, a structural gene, an UTR or as a

PI 3' termination sequence. They are also useful as tools for genetic

PI mapping, and identification of a particular individual plant or for

PI clustering a group of plants with a common trait. AAA78433 to AAA78630

PI and AAB24605 to AAB25099 represent the specifically claimed

PI sequences and/or sequences of genomic DNA encompassing complete genes,

PI portions of genes, and/or intergenic regions, collectively referred to

PI as sequence-determined DNA fragments (SDFs), from corn plants and

PI Arabidopsis thaliana. The SDFs are promoters, structural genes,

PI untranslated regions (UTRs), or 3' termination sequences. They can be

PI used for expressing a gene product and controlling expression of a

PI target gene, either as a promoter, a structural gene, an UTR or as a

PI 3' termination sequence. They are also useful as tools for genetic

PI mapping, and identification of a particular individual plant or for

PI clustering a group of plants with a common trait. AAA78433 to AAA78630

PI and AAB24605 to AAB25099 represent the specifically claimed

PI sequences and/or sequences of genomic DNA encompassing complete genes,

PI portions of genes, and/or intergenic regions, collectively referred to

PI as sequence-determined DNA fragments (SDFs), from corn plants and

PI Arabidopsis thaliana. The SDFs are promoters, structural genes,

PI untranslated regions (UTRs), or 3' termination sequences. They can be

PI used for expressing a gene product and controlling expression of a

PI target gene, either as a promoter, a structural gene, an UTR or as a

PI 3' termination sequence. They are also useful as tools for genetic

PI mapping, and identification of a particular individual plant or for

PI clustering a group of plants with a common trait. AAA78433 to AAA78630

PI and AAB24605 to AAB25099 represent the specifically claimed

PI sequences and/or sequences of genomic DNA encompassing complete genes,

PI portions of genes, and/or intergenic regions, collectively referred to

PI as sequence-determined DNA fragments (SDFs), from corn plants and

PI Arabidopsis thaliana. The SDFs are promoters, structural genes,

PI untranslated regions (UTRs), or 3' termination sequences. They can be

PI used for expressing a gene product and controlling expression of a

PI target gene, either as a promoter, a structural gene, an UTR or as a

PI 3' termination sequence. They are also useful as tools for genetic

PI mapping, and identification of a particular individual plant or for

PI clustering a group of plants with a common trait. AAA78433 to AAA78630

PI and AAB24605 to AAB25099 represent the specifically claimed

PI sequences and/or sequences of genomic DNA encompassing complete genes,

PI portions of genes, and/or intergenic regions, collectively referred to

PI as sequence-determined DNA fragments (SDFs), from corn plants and

PI Arabidopsis thaliana. The SDFs are promoters, structural genes,

PI untranslated regions (UTRs), or 3' termination sequences. They can be

PI used for expressing a gene product and controlling expression of a

PI target gene, either as a promoter, a structural gene, an UTR or as a

PI 3' termination sequence. They are also useful as tools for genetic

PI mapping, and identification of a particular individual plant or for

PI clustering a group of plants with a common trait. AAA78433 to AAA78630

PI and AAB24605 to AAB25099 represent the specifically claimed

PI sequences and/or sequences of genomic DNA encompassing complete genes,

Query Match 20.1%; Score 68; DB 21; Length 78;  
 Best Local Similarity 30.1%; Pred. NO. 1.3;  
 Matches 22; Conservative 9; Mismatches 26; Indels 16; Gaps 4;  
 QY 2 KSIAIIFIVLVAFCILEGIV-EAGFGCPFN-----AGKCHRHCKSIRRG-G 47  
 DB 4 kllqvstvmiiftllvgvvaneglgkpkqkneilkqsnvcvaacdsmcvkkrkgag 63  
 QY 48 FCRGTFRRTTCVY 60  
 DB 64 ycspskk--cycy 74

Search completed: September 17, 2002, 15:44:15  
 Job time: 126 sec



OC Buthoidea; Buthidae; Leiurus.

```

OX NCBI_TaxID=6884;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=93326112; PubMed=8333834;
RA Cociancich S., Goyffon M., Bontems F., Bulet P., Bouet F., Menez A.,
RA Hoffmann J.A.;
RT "Purification and characterization of a scorpion defensin, a 4kDa
RT antibacterial peptide presenting structural similarities with insect
RT defensins and scorpion toxins.";
RL Biochem. Biophys. Res. Commun. 194:17-22(1993).
CC -!- FUNCTION: ANTIBACTERIAL PROTEIN AGAINST GRAM-POSITIVE BACTERIA;
CC MAY ACT VIA MEMBRANE-PERMEABILISATION OF THESE CELLS.
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR PIR; JN0613; JN0613.
DR HSP; P10891; 1ICA.
DR InterPro; IPR001542; Arthro.defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Antibiotic.
FT DISULFID 4 25 BY SIMILARITY.
FT DISULFID 11 33 BY SIMILARITY.
FT DISULFID 15 35 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4326 MW; DF35FB21EC83FB84 CRC64;

Query Match 52.2%; Score 177; DB 1; Length 38;
Best Local Similarity 73.0%; Pred. No. 1.4e-13;
Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

-QY 25 GFGCPFNAGKCHRRCKSIRRRGGFCGRGTRTTCVCYR 61
Db 1 GFGCPNAGACHRRCKSIRRRGGYCGAGFKQTCYR 37

-RESULT 3
ID DEFLAESC STANDARD; PRT; 38 AA.
AC P80154;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE Defensin.
OS Aeschna cyanea (Dragonfly) (Blue darner).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Palaeoptera; Odonata; Anisoptera; Aeshnidae; Aeschna.
OX NCBI_TaxID=12921;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=93049356; PubMed=1425705;
RA Bulet P., Cociancich S., Reuland M., Sauber F., Bischoff R., Hegy G.,
RA van Dorsseleer A., Hetru C., Hoffmann J.A.;
RT "A novel insect defensin mediates the inducible antibacterial
RT activity in larvae of the dragonfly Aeschna cyanea (Paleoptera,
RT Odonata).";
RL Eur. J. Biochem. 209:977-984(1992).
CC -!- FUNCTION: MEDIATES THE INDUCIBLE ANTIBACTERIAL ACTIVITY IN LARVAE
CC OF A.CYANEA.
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR PIR; S27242; S27242.
DR InterPro; IPR001542; Arthro.defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Insect immunity; Antibiotic.
FT DISULFID 4 26 BY SIMILARITY.
FT DISULFID 11 34 BY SIMILARITY.
FT DISULFID 15 36 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4180 MW; 7C884A8FDF54A46A CRC64;

Query Match 39.4%; Score 133.5; DB 1; Length 38;
Best Local Similarity 55.3%; Pred. No. 9.7e-09;

```

```

Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCPFNAGKCHRRCKSIRRRGGFCGRGTRTTCVCYR 61
Db 1 GFGCPNAGACHRRCKSIRRRGGYCGAGFKQTCYR 38

RESULT 4
ID DEFLMYTGA STANDARD; PRT; 38 AA.
AC P80571;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Defensin MGD-1.
GN PH3.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=97025339; PubMed=8925841;
RA Hubert F., Noel T., Roch P.;
RT "A member of the arthropod defensin family from edible Mediterranean
RT mussels (Mytilus galloprovincialis).";
RL Eur. J. Biochem. 240:302-306(1996).
RN [2]
RP ERRATUM.
RA Hubert F., Noel T., Roch P.;
RA Eur. J. Biochem. 240:815-815(1996).
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR InterPro; IPR001542; Arthro.defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Antibiotic.
FT DISULFID 4 25 BY SIMILARITY.
FT DISULFID 10 33 BY SIMILARITY.
FT DISULFID 14 35 BY SIMILARITY.
FT DISULFID 21 38 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4275 MW; 6F1C0CDCF0E69E76 CRC64;

Query Match 37.8%; Score 128; DB 1; Length 38;
Best Local Similarity 63.2%; Pred. No. 4e-08;
Matches 24; Conservative 2; Mismatches 10; Indels 2; Gaps 2;

QY 25 GFGCPFNAGKCHRRCKSIRRRGGFCGRGTRTTCVCYR 61
Db 1 GFGCPNAGACHRRCKSIRRRGGYCGAGFKQTCYR 37

RESULT 5
ID DEFLMYTED STANDARD; PRT; 37 AA.
AC P81610;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Defensin A.
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=Blood;
RX MEDLINE=96355569; PubMed=8702979;
RA Charlet M., Chernysh S., Philippe H., Hetru C., Hoffmann J.A.,
RA Bulet P.;
RT "Innate immunity. Isolation of several cysteine-rich antimicrobial
RT peptides from the blood of a mollusc, Mytilus edulis.";

```

[illegible]

DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-MAR-1992 (Rel. 21, Last annotation update)  
DE Defensin, isoforms B and C.  
OS Zophobas atratus.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
OC Cucujiformia; Tenebrionidae; Zophobas.  
OX NCBI\_TaxID=7074;  
RN [1]  
RN RN  
RP SEQUENCE.  
RC TISSUE=Hemolymph;  
RX MEDLINE=92105112; PubMed=1761552;  
RA Boulet P., Coclicanich S., Dimarcq J.-L., Lambert J., Reichhart J.-M.,  
RA Hoffmann D., Hetru C., Hoffmann J.A.;  
RA "Insect immunity. Isolation from a coleopteran insect of a novel  
RT inducible antibacterial peptide and of new members of the insect  
RT defensin family.";   
RL J. Biol. Chem. 266:24520-24525(1991).  
RL -!- FUNCTION: INVOLVED IN ANTI GRAM-POSITIVE ACTIVITY OF IMMUNE  
CC HEMOLYMPH OF Z.ATRATUS.  
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
DR PIR; B41711; B41711.  
DR PIR; C41711; C41711.  
DR HSP; P10891; 1ICA.  
DR InterPro; IPR001542; Arthro\_defensin.  
DR InterPro; IPR003614; Knot1.  
DR Pfam; PF01097; Arthro\_Defensin; 1.  
DR PRINTS; PR00271; DEFENSIN.  
DR SMART; SM00505; Knot1; 1.  
DR PROSITE; PS00425; ARTHROPOD\_DEFENSINS; 1.  
KW Insect immunity; Antibiotic.  
FT DISULFID 3 34 BY SIMILARITY.  
FT DISULFID 20 40 BY SIMILARITY.  
FT DISULFID 24 42 BY SIMILARITY.  
FT VARIANT 30 30 R -> T (IN ISOFORM C).  
SQ SEQUENCE 43 AA; 4453 MW; D63E7B100987FD2F CRC64;  
Query Match 23.7%; Score 80.5; DB 1; Length 43;  
Best Local Similarity 39.5%; Pred. No. 0.0088;  
Matches 15; Conservative 6; Mismatches 12; Indels 5; Gaps 2;  
QY 25 GF---GCPFNACKCHRHCKSIIRRGFGCGTFRTCVC 59  
|| | : | : || : ||| : | : ||  
DB 7 GFEIAGTLNSAACGHCLALGRGGYCNS--KSCVC 42  
RESULT 9  
ID SCX2 ANDAU STANDARD; PRT; 85 AA.  
AC P01484;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neurotoxin II precursor (Aah II).  
OS Androctonus australis hector (Sahara scorpion).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Scorpiones;  
OC Buthoidea; Buthidae; Androctonus.  
OX NCBI\_TaxID=70175;  
RN [1]  
RN RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90037062; PubMed=2808423;  
RA Bougis P.E., Rochat H., Smith L.A.;  
RA "precursors of Androctonus australis scorpion neurotoxins. Structures  
RT of precursors, processing outcomes, and expression of a functional  
RT recombinant toxin II".  
RL J. Biol. Chem. 264:19259-19265(1989).  
RN [2]  
RN RN  
RP SEQUENCE OF 20-82.  
RX MEDLINE=73025153; PubMed=4342910;  
RA Rochat H., Rochat C., Sampieri F., Miranda F., Liassitzky S.;  
RA "the amino-acid sequence of neurotoxin II of Androctonus australis

DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-MAR-1992 (Rel. 21, Last annotation update)  
DE Defensin, isoforms B and C.  
OS Zophobas atratus.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
OC Cucujiformia; Tenebrionidae; Zophobas.  
OX NCBI\_TaxID=7074;  
RN [1]  
RN RN  
RP SEQUENCE.  
RC TISSUE=Hemolymph;  
RX MEDLINE=92105112; PubMed=1761552;  
RA Boulet P., Coclicanich S., Dimarcq J.-L., Lambert J., Reichhart J.-M.,  
RA Hoffmann D., Hetru C., Hoffmann J.A.;  
RT "Insect immunity. Isolation from a coleopteran insect of a novel  
RT inducible antibacterial peptide and of new members of the insect  
RT defensin family."; J. Biol. Chem. 266:24520-24525(1991).  
RL J. Biol. Chem. 266:24520-24525(1991).  
RR CC  
CC -!- FUNCTION: INVOLVED IN ANTI GRAM-POSITIVE ACTIVITY OF IMMUNE  
CC HEMOLYMPH OF Z.ATRATUS.  
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
DR PIR; B41711; B41711.  
DR PIR; C41711; C41711.  
DR HSP; P10891; 1ICA.  
DR InterPro; IPR001542; Arthro\_defensin.  
DR InterPro; IPR003614; Knot1.  
DR Pfam; PF01097; Arthro\_Defensin; 1.  
DR PRINTS; PR00271; DEFENSIN.  
DR SMART; SM00505; Knot1; 1.  
DR PROSITE; PS00425; ARTHROPOD\_DEFENSINS; 1.  
DR Insect immunity; Antibiotic.  
KW DISULFID 3 34 BY SIMILARITY.  
FT DISULFID 20 40 BY SIMILARITY.  
FT DISULFID 24 42 BY SIMILARITY.  
FT VARIANT 30 30 R -> T (IN ISOFORM C).  
SQ SEQUENCE 43 AA; 4453 MW; D63E7B100987FD2F CRC64;  
Query Match 23.7%; Score 80.5; DB 1; Length 43;  
Best Local Similarity 39.5%; Pred. No. 0.0088;  
Matches 15; Conservative 6; Mismatches 12; Indels 5; Gaps 2;  
QY 25 GF---GCPFNACKCHRHCKSIERRGGFCRGTRTCVC 59  
|| | : | : || : ||||| : : |||  
DB 7 GFEIAGTLNSAACGHCLALGRGGYCNS--KSCVC 42  
RESULT 9  
ID SCX2 ANDAU STANDARD; PRT; 85 AA.  
AC P01484;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neurotoxin II precursor (Aah II).  
OS Androctonus australis hector (Sahara scorpion).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Scorpiones;  
OC Buthoidea; Buthidae; Androctonus.  
OX NCBI\_TaxID=70175;  
RN [1]  
RN RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90037062; PubMed=2808423;  
RA Bougis P.E., Rochat H., Smith L.A.;  
RT "precursors of Androctonus australis scorpion neurotoxins. Structures  
RT of precursors, processing outcomes, and expression of a functional  
RT recombinant toxin II."; J. Biol. Chem. 264:19259-19265(1989).  
RL J. Biol. Chem. 264:19259-19265(1989).  
RN [2]  
RN RN  
RP SEQUENCE OF 20-82.  
RX MEDLINE=73025153; PubMed=4342910;  
RA Rochat H., Rochat C., Sampieri F., Miranda F., Liassitzky S.;  
RT "The amino-acid sequence of neurotoxin II of Androctonus australis

[illegible]

```

CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN FLOWERS AND IN YOUNG FRUITS.
CC -!- DEVELOPMENTAL STAGE: ACCUMULATE DURING RIPENING.
CC -!- SIMILARITY: BELONGS TO THE GAMMA-PUROTHIONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95730; CAA65046.1; -.
DR HSSP; P20230; IGPT.
DR Mendel; 29862; Capan;1533;29862.
DR InterPro; IPR002118; Gamma-thionin.
DR Pfam; IPR003614; Gamma-thionin.
DR ProDom; PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knot1; 1.
DR PROSITE; PS00940; GAMMA_THIONIN; 1.
KW Defensin; Fungicide; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 74 DEFENSIN J1-2.
FT DISULFID 30 74 BY SIMILARITY.
FT DISULFID 41 61 BY SIMILARITY.
FT DISULFID 47 68 BY SIMILARITY.
FT DISULFID 51 70 BY SIMILARITY.
SQ SEQUENCE 74 AA; 8249 MW; D92D8F06F39E1552 CRC64;

Query Match 21.1%; Score 71.5; DB 1; Length 74;
Best Local Similarity 30.0%; Pred. No. 0.13;
Matches 21; Conservative 8; Mismatches 28; Indels 13; Gaps 3;

QY 2 KSAIIFVLVAFCLIEDGIVEAGFGCP-----FNAGKCHRHCKSIIRRGGFCRG 51
DB 6 KVIATIFLMM--LVFATGWAEATCESQSHRFKGLCFSGNCSGVSCHTGFNGGHCGRG 63
QY 52 TPTTCVCVYR 61
DB 64 -FRRRCFCYR 72

RESULT 13
DEFL_TENMO
ID DEFL_TENMO STANDARD; PRT; 84 AA.
AC Q27023;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tenebrion 1 precursor.
OS Tenebrio molitor (yellow mealworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 49-59, AND DISULFIDE BONDS.
RC TISSUE-Larval hemolymph;
RX MEDLINE=95096025; PubMed=7798186;
RA Moon H.J., Lee S.Y., Kurata S., Natori S., Lee B.L.;
RT "Purification and molecular cloning of cDNA for an inducible
RT antibacterial protein from larvae of the coleopteran, Tenebrio
RT molitor.";
RL J. Biochem. 116:53-58(1994).
CC -!- FUNCTION: BACTERICIDAL PROTEIN PRODUCED IN RESPONSE TO INJURY. IT
CC IS CYTOTOXIC TO GRAM-POSITIVE BACTERIA.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17670; BAA04552.1; -.
DR HSSP; P10891; LICA.
DR InterPro; IPR001542; Arthro_defensin.
DR Pfam; PF01097; Arthro_defensin; 1.
DR PRINTS; PR00271; DEFENSIN.
DR SMART; SM00505; Knot1; 1.
DR PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.
KW Insect immunity; Antibiotic; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 41
FT CHAIN 42 84 TENECIN 1.
FT DISULFID 44 75
FT DISULFID 61 81
FT DISULFID 65 83
SQ SEQUENCE 84 AA; 9176 MW; 0367C5070468BE60 CRC64;

Query Match 21.1%; Score 71.5; DB 1; Length 84;
Best Local Similarity 27.3%; Pred. No. 0.15;
Matches 21; Conservative 7; Mismatches 26; Indels 23; Gaps 3;

QY 4 IAIIFIVLVAFCILEDGI---VEAG-----FGCFPNAGKCHRHCKSI 42
DB 9 VACFFILQIAAFPLEEAATAETEGSEHVRKRVCTDILSVEAKGVKLNDAACAAHCLPR 68
QY 43 RRRGGFCRGFTTCVC 59
DB 69 GRGGYCNG--KRVVCV 83

RESULT 14
DEFC_AEDAE
ID DEFC_AEDAE STANDARD; PRT; 99 AA.
AC P81603; Q9Y0F1; Q9Y0F0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Defensin C precursor.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Aedes.
OX NCBI_TaxID=7159;
RN [1]
RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, INDUCTION, AND SUBCELLULAR
RP LOCATION.
RC STRAIN=LIVERPOOL; TISSUE=Fat body;
RX MEDLINE=99124369; PubMed=9927179;
RA Lowenberger C.A., Smartt C.T., Bulet P., Ferdig M.T., Severson D.W.,
RA Hoffmann J.A., Christensen B.M.;
RT "Insect immunity: molecular cloning, expression, and characterization
RT of cDNAs and genomic DNA encoding three isoforms of insect defensin in
RT Aedes aegypti.";
RL Insect Mol. Biol. 8:107-118(1999).
RN [2]
RP SEQUENCE OF 60-99, AND INDUCTION.
RC STRAIN=LIVERPOOL;
RX MEDLINE=95360030; PubMed=7633471;
RA Lowenberger C., Bulet P., Charlet M., Hetru C., Hodgeman B.,
RA Christensen B.M., Hoffmann J.A.;
RT "Insect immunity: isolation of three novel inducible antibacterial
RT defensins from the vector mosquito, Aedes aegypti.";
RL Insect Biochem. Mol. Biol. 25:867-873(1995).
RN [3]

```







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:42:49 ; Search time 18.35 seconds  
(without alignments)  
319.425 million cell updates/sec

Title: US-09-829-481-4  
Perfect score: 339  
Sequence: 1 MKSIAIFIVLVAFCILEDG.....IRRRGFCRGFTTTCVCYR 61

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	177	52.2	38	JN0613	defensin 4K - scor
2	133.5	39.4	38	S27242	defensin - blue da
3	128	37.8	38	S74088	defensin - Mediter
4	80.5	23.7	43	B41711	defensin B - beetl
5	79.5	23.5	85	NTSR2A	neurotoxin II prec
6	74.5	22.0	43	C41711	defensin C - beetl
7	72.5	21.4	43	JC2554	holotricin I - Hol
8	71.5	21.1	84	JX0332	tenecin I precursor
9	69.5	20.5	65	NTSR3B	neurotoxin XI - sc
10	69	20.4	75	S11156	psa10 protein - c
11	68.5	20.2	72	T06599	disease resistance
12	68	20.1	94	A31792	sapacin precursor
13	67.5	19.9	79	T06381	proteinase inhibit
14	66.5	19.6	64	NTSR3B	neurotoxin III - s
15	66.5	19.6	86	JN0671	Nat-channel-blocki
16	66	19.5	94	S12558	defensin A precurs
17	65.5	19.3	64	NTSR5L	neurotoxin V - Egy
18	65.5	19.3	64	NTSR5M	neurotoxin V - sco
19	64.5	19.0	798	A40526	integrin beta-7 ch
20	63.5	18.7	77	B84433	proteinase inhibit
21	63.5	18.7	80	T10823	antifungal protein
22	63.5	18.7	768	B41029	integrin beta-8 ch
23	63.5	18.7	769	A41029	integrin beta-8 ch
24	63	18.6	90	B84867	probable trypsin i
25	62	18.3	990	I51618	nucleolar phosphop
26	61.5	18.1	77	S30578	protease inhibitor
27	61.5	18.1	80	T10176	antifungal protein
28	61.5	18.1	154	T17816	hypothetical prote
29	61	18.0	40	JU0225	sapacin C - flesh

30 60.5 17.8 80 2 F96787 protein T4O12.7 [i  
31 60.5 17.8 126 2 T21762 hypothetical prote  
32 60.5 17.8 369 2 C75378 glutamate racemase  
33 60.5 17.8 806 2 A46271 integrin beta-7 ch  
34 60 17.7 79 2 T10243 antifungal protein  
35 60 17.7 84 2 S62868 toxin gamma precu  
36 59 17.4 84 2 S21158 neurotoxin TsvII p  
37 58.5 17.3 75 2 S51637 probable proteinas  
38 58.5 17.3 77 2 T14395 proteinase inhibit  
39 58.5 17.3 80 2 T10183 antifungal protein  
40 58.5 17.3 80 2 F96591 probable antifunga  
41 58.5 17.3 640 2 T45924 protein kinase-lik  
42 58 17.1 304 2 F69230 pyruvate formate-1  
43 58 17.1 318 2 E87929 protein T22H2.6 [i  
44 58 17.1 358 2 T25137 hypothetical prote  
45 58 17.1 964 2 JC5545 integrin beta-4 pr

#### ALIGNMENTS

##### RESULT 1

JN0613

defensin 4K - scorpion (Leiurus quinquestriatus)

N:Alternate names: antibacterial 4K peptide

C:Species: Leiurus quinquestriatus hebraeus

C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 03-Feb-1994

C:Accession: JN0613

R:Cociancich, S.; Goyffon, M.; Bontems, F.; Bulet, P.; Menez, A.; Hoffmann

Biochem. Biophys. Res. Commun. 194, 17-22, 1993

A:Title: Purification and characterization of a scorpion defensin, a 4kDa antibacteri

A:Reference number: JN0613; MUID:93326112

A:Accession: JN0613

A:Molecule type: protein

A:Residues: 1-38 <COC>

A>Note: this protein is similar to scorpion toxins and insect defensins

Query Match 52.2%; Score 177; DB 2; Length 38;

Best Local Similarity 73.0%; Pred. No. 5.4e-13;

Matches 27; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 25 GFGCFPNAGKCHRHCKSIRRRGFCRGFTTTCVCYR 61

||||| | ||||| : ||||| : |||||

Db 1 GFGCPLNQACHRHCKSIRRRGGYCAAGFKQCTCYR 37

##### RESULT 2

S27242

defensin - blue darner

C:Species: Aeschna cyanea (blue darner)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-1999

C:Accession: S27242

R:Bulet, P.; Cociancich, S.; Reuland, M.; Sauber, F.; Bischoff, R.; Hegy, G.; van Dor

Eur. J. Biochem. 209, 977-984, 1992

A:Title: A novel insect defensin mediates the inducible antibacterial activity in lar

A:Reference number: S27242; MUID:93049356

A:Accession: S27242

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-38 <BUL>

Query Match 39.4%; Score 133.5; DB 2; Length 38;

Best Local Similarity 55.3%; Pred. No. 3.6e-08;

Matches 21; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 25 GFGCFPNAGKCHRHCKSIR-RRGFCRGFTTTCVCYR 61

||||| : ||||| : ||||| : ||||| : |||||

Db 1 GFGCPLDQMOCHRHCKOTTGRRGGYCSGLKUTCTCYR 38

##### RESULT 3

S74088  
defensin - Mediterranean mussel  
C:Species: Mytilus galloprovincialis (Mediterranean mussel)  
C:Date: 11-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 04-Feb-2000  
C:Accession: S74088  
R:Hubert, F.; Noel, T.; Roch, P.  
Eur. J. Biochem. 240, 302-306, 1996  
A:Title: A member of the arthropod defensin family from edible Mediterranean mussels (Mytilus galloprovincialis)  
A:Reference number: S74088; MUID:97025339  
A:Accession: S74088  
A:Molecule type: protein  
A:Residues: 1-38 <HUB>  
A:Experimental source: hemolymph  
C:Keywords: antibiotic  
F:4-25,10-33,14-35,21-38/Disulfide bonds: #status predicted

Query Match 37.8%; Score 128; DB 2; Length 38;  
Best Local Similarity 63.2%; Pred.No.1.5e-07;  
Matches 24; Conservative 2; Mismatches 10; Indels 2; Gaps 2;

OY 25 GFGCFNAGCKHRHCKSIKRR-GGFCRGTFRTTCVCR 61  
||||| : ||||| ||| : ||| : |||  
Db 1 GFGCP-NNYOCHRRCKSIKRRGGYCGGXHRLRCTCYR 37

RESULT 4  
B41711  
defensin B - beetle (Zophobas atratus)  
C:Species: Zophobas atratus  
C:Date: 24-Jul-1992 #sequence\_revision 24-Jul-1992 #text\_change 03-May-1996  
C:Accession: B41711  
R:Bulet, P.; Cociancich, S.; Dimarcq, J.L.; Lambert, J.; Reichhart, J.M.; Hoffmann, D.; J. Biol. Chem. 266, 24520-24525, 1991  
A:Title: Insect immunity. Isolation from a coleopteran insect of a novel inducible anti-insect defensin B from the beetle Zophobas atratus  
A:Reference number: A41711; MUID:92105112  
A:Accession: B41711  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-43 <BUL>  
C:Superfamily: sapecin precursor

Query Match 23.7%; Score 80.5; DB 2; Length 43;  
Best Local Similarity 39.5%; Pred.No.0.03;  
Matches 15; Conservative 6; Mismatches 12; Indels 5; Gaps 2;

OY 25 GF---GCFNAGCKHRHCKSIKRRGGFCRGTFRTTCVC 59  
||| : ||| : ||| : ||| : |||  
Db 7 GFETAGTKLSAAGAHCLALGRGGYCNLS--KSVCCV 42

RESULT 5  
NTR2A  
neurotoxin II precursor [validated] - Sahara scorpion  
N:Alternate names: mammalian neurotoxin (clone pcb-402)  
C:Species: Androctonus australis (Sahara scorpion)  
C:Date: 24-Apr-1984 #sequence\_revision 12-Apr-1996 #text\_change 15-Sep-2000  
C:Accession: D34444; A01744  
R:Bougis, P.E.; Rochat, H.; Smith, L.A. J. Biol. Chem. 264, 19259-19265, 1989  
A:Title: Precursors of Androctonus australis scorpion neurotoxins. Structures of precursors of neurotoxin II and neurotoxin I  
A:Reference number: A34444; MUID:90037062  
A:Accession: D34444  
A:Molecule type: mRNA  
A:Residues: 1-85 <BOU>  
A:Cross-references: GB:M27704; NID:gi61144; PIDN:AAA29949.1; PID:gi61145; GB:J05102  
R:Experimental source: clone pcb-402  
R:Rochat, H.; Rochat, C.; Sampieri, F.; Miranda, F.; Lissitzky, S. Eur. J. Biochem. 28, 381-388, 1972  
A:Title: The amino-acid sequence of neurotoxin II of Androctonus australis Hector.  
A:Reference number: A91198; MUID:73025153  
A:Accession: A01744

```
A:Molecule type: protein
A:Residues: 20-83 <ROC>
R:Kopeyan, C.; Martinec, G.; Lissitzky, S.; Miranda, F.; Rochat, H. Eur. J. Biochem. 47, 483-489, 1974
A>Title: Disulfide bonds of toxin II of the scorpion Androctonus australis Hector.
A:Reference number: A91225; MUID:75057756
A:Contents: annotation; disulfide bonds
R:Housset, D.; Habersetzer-Rochat, C.; Astier, J.P.; Fontecilla-Camps, J.C. J. Mol. Biol. 238, 88-103, 1994
A>Title: Crystal structure of toxin II from the scorpion Androctonus australis Hector
A:Reference number: A58108; MUID:94194522
A:Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83
R:Fontecilla-Camps, J.C.; Housset, D. submitted to the Brookhaven Protein Data Bank, September 1994
A:Reference number: A67142; PDB:LPTX
A:Contents: annotation; X-ray crystallography, 1.3 angstroms, residues 20-83
C:Superfamily: scorpion neurotoxin
C:Keywords: amidated carboxyl end; neurotoxin; venom
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-83/Product: neurotoxin II #status experimental <MAT>
F:31-82;35-55;41-65;45-67/Disulfide bonds: #status experimental
F:83/Modified site: amidated carboxyl end (His) (amide in mature form from following
```

---

```
Query Match          23.5%   Score 79.5; DB 1; Length 85;
Best Local Similarity 24.2%; Pred. No. 0.062;
Matches 16; Conservative 16; Mismatches 27; Indels 7; Gaps 3;
```

---

```
QY    1 MKSIAIPIVLVAPCIILEDGIVEAGFCPPFNAGK--CHRHCKSIIRRGGFGR--GTFTT 55
      I::L::I::L::I::L::I::L::I::L::I::L::I::L::I::L::I::L::I::L::I::L::
Db     6 MISLALELVTVGE--SVRKDGYIVDVNCTYPCGRNAYCNBECTKLKGESGYCWASPYGN 63

QY    56 TCVCYR 61
      |::|::
Db     64 ACYCVR 69
```

---

```
RESULT           6
C41711
defensin C - beetle (Zophobas atratus)
```

---

```
C:Species: Zophobas atratus
C>Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 03-May-1996
C:Accession: C41711
R:Bulet, P.; Cociancich, S.; Dimarcq, J.L.; Lambert, J.; Reichhart, J.M.; Hoffmann, D. J. Biol. Chem. 266, 24520-24525, 1991
A>Title: Insect immunity. Isolation from a coleopteran insect of a novel inducible alpha-A:Reference number: A41711; MUID:92105112
A:Accession: C41711
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-43 <BUH>
C:Superfamily: sapecin precursor
```

---

```
Query Match          22.0%   Score 74.5; DB 2; Length 43;
Best Local Similarity 36.8%; Pred. NO. 0.14;
Matches 14; Conservative 6; Mismatches 13; Indels 5; Gaps 2;
```

---

```
QY    25 GF---GCFFNAGKHRRICKSTIRRGGFGCGFRFTTCVC 59
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     7 GFETATIGIKLSACGAHLALGRTGGYCNS--KSVCVC 42
```

---

```
RESULT           7
JC2554
holotricin 1 - Holotrachia diomphalia
```

---

```
C:Species: Holotrachia diomphalia
C>Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 17-Mar-1999
C:Accession: JC2554
R:Lee, S.Y.; Moon, H.J.; Kawabata, S.; Kurata, S.; Natori, S.; Lee, B.I. Biol. Pharm. Bull. 18, 457-459, 1995
A>Title: A sapecin homologue of Holotrachia diomphalia: Purification, sequencing and A:Reference number: JC2554; MUID:96031330
```

A:Accession: JC2554

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-43 <LEE>

C:Superfamily: sapecin precursor

Query Match 21.4%; Score 72.5; DB 2; Length 43;

Best Local Similarity 39.4%; Pred. No. 0.23; Mismatches 5; Gaps 1;

Matches 13; Conservative 5; Indels 3; Gaps 3;

QY 27 GCFPNAGKCHRHCKSIIRRRGGFCGRTTCVC 59

Db 12 GIAINDSACAHCCLAMRRKGSGCK---QGVVC 41

RESULT 8

JX0332

tenecin 1 precursor - yellow mealworm

C:Species: Tenebrio molitor (yellow mealworm)

C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 20-Jun-2000

C:Accession: JX0332

R:Moore, H.J.; Lee, S.Y.; Kurata, S.; Natori, S.; Lee, B.L.

J. Biochem. 116, 53-58, 1994

A:Title: Purification and molecular cloning of cDNA for an inducible antibacterial prote

A:Reference number: JX0332; MUID:95096025

A:Accession: JX0332

A:Molecule type: mRNA

A:Residues: 1-84 <MOO>

A:Cross-references: GB:D17670; NID:g1235940; PIDN:BA04552.1; PID:g1235941

A:Experimental source: larva

C:Comment: Tenecin 1 is an antibacterial protein induced in the hemolymph of larvae of t

C:Superfamily: sapecin precursor

F:1-41/Domain: signal sequence #status predicted <SIG>

F:42-84/Product: tenecin 1 #status predicted <MAT>

F:44-75,61-81,65-83/Disulfide bonds: #status predicted

Query Match 21.1%; Score 71.5; DB 2; Length 84;

Best Local Similarity 27.3%; Pred. No. 0.48; Mismatches 7; Indels 26; Gaps 3;

Matches 21; Conservative 7; Indels 23; Gaps 3;

QY 4 IAIIFIVLVAFCILEDGI---VEAG-----FGCFPNAGKCHRHCKSI 42

Db 9 VACFFIIQIAAFPLEEATAAEIEQGEHIRVKRTCDILSVEAKGVKLNDAACAAHCLFR 68

QY 43 RRRGGFCGRTTCVC 59

Db 69 GRSGGYCNG--KRVCVC 83

RESULT 9

NTSREB

neurotoxin XI - scorpion (Buthus occitanus)

C:Species: Buthus occitanus tunetanus

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 23-Aug-1996

C:Accession: A01746

R:Martin, M.F.; Rochat, H.

Toxicon 22, 279-291, 1984

A:Title: Purification of thirteen toxins active on mice from the venom of the North Afric

A:Reference number: A94316; MUID:84224814

A:Accession: A01746

A:Molecule type: protein

A:Residues: 1-65 <MAR>

C:Superfamily: scorpion neurotoxin

C:Keywords: neurotoxin; venom

F:12-63,16-36,22-46,26-48/Disulfide bonds: #status predicted

Query Match 20.5%; Score 69.5; DB 1; Length 65;

Best Local Similarity 26.0%; Pred. No. 0.66; Mismatches 10; Indels 22; Gaps 2;

Matches 13; Conservative 10; Indels 5; Gaps 2;

QY 17 LEDGIVEAGFCGCPFNAGK---CHRHCKSIIRRRGGFCR--GTFRTTCVCYR 61

Db 1 LKDGIVDDRNCTYFCGTNAYCNEECYKLGESYCCQWVGRYGNACWCYK 50

RESULT 10

S11156

PSA10 protein - cowpea

C:Species: Vigna unguiculata (cowpea)

C:Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jun-2000

C:Accession: S11156

R:Ishibashi, N.; Yamauchi, D.; Minamikawa, T.

Plant Mol. Biol. 15, 59-64, 1990

A:Title: Stored mRNA in cotyledons of Vigna unguiculata seeds: nucleotide sequence of

A:Reference number: S11156; MUID:91355865

A:Accession: S11156

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-75 <ISH>

A:Cross-references: EMBL:X16877; NID:g22075; PID:g22076

C:Superfamily: gamma-thionin

Query Match 20.4%; Score 69; DB 2; Length 75;

Best Local Similarity 33.3%; Pred. No. 0.84; Mismatches 6; Indels 14; Gaps 4;

Matches 24; Conservative 6; Indels 14; Gaps 4;

QY 2 KSTIA-IIFIVLVAFCILE-----DGIVEAGFCGCPFNAGKCHRHCKSIIR- RGGFC 49

Db 4 KSTAGLCFLFLVFAQEVVVOSEAKTCENLVDTYRGCFCTGSCDDHCKNKEHLLSRC 63

QY 50 RGTFRITTCVCYR 61

Db 64 RDDVR--CWCTR 73

RESULT 11

T06599

disease resistance response protein 230 - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jun-2000

C:Accession: T06599

R:Chiang, C.C.; Hadwiger, L.A.

Mol. Plant Microbe Interact. 4, 324-331, 1991

A:Title: The Fusarium solani-induced expression of a pea gene family encoding high cy

A:Reference number: Z15787; MUID:92190628

A:Accession: T06599

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-72 <CHI>

A:Cross-references: EMBL:L01578; NID:gi169073; PIDN:AAA79117.1; PID:gl169074

A:Experimental source: cv. Alaska

C:Genetics:

A:Gene: DRR230-a

C:Superfamily: gamma-thionin

Query Match 20.2%; Score 68.5; DB 2; Length 72;

Best Local Similarity 28.4%; Pred. No. 0.92; Mismatches 11; Indels 13; Gaps 3;

Matches 19; Conservative 11; Indels 13; Gaps 3;

QY 2 KSTAIIFIVLVAFCILEDGIVEAGFCGCPFNAGK-----CHRHCKSIIRRG---GFCR 50

Db 4 KSLACLFLLLVFAQEIWVSEANTCENLAGSKYGVCGGCDRHCR--QEGATSGRCR 61

QY 51 GTFRTTC 57

Db 62 DDFRCWC 68

RESULT 12

A31792

sapecin precursor - flesh fly (Sarcophaga peregrina)

2

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 17, 2002, 15:43:24 ; Search time 25.63 seconds  
(without alignments)  
411.732 Million cell updates/sec

Title: US-09-829-481-4  
Perfect score: 339  
Sequence: 1 MKSIAIFIVLVAFCILEDG.....IRRGGFCRGTRFTTCVCYR 61

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19.\*  
1: sp-archaea.\*  
2: sp-bacteria.\*  
3: sp-fungi.\*  
4: sp-human.\*  
5: sp-invertebrate.\*  
6: sp-mammal.\*  
7: sp-mhc.\*  
8: sp-organelle.\*  
9: sp-phase.\*  
10: sp-plant.\*  
11: sp-rodent.\*  
12: sp-virus.\*  
13: sp-vertebrate.\*  
14: sp-unclassified.\*  
15: sp-rvirus.\*  
16: sp-bacteriap.\*  
17: sp-archaeap.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	154.5	45.6	81	5 Q9Y0A9	Q9Y0A9 mytilus gal
2	152	44.8	82	5 Q9U6U0	Q9U6U0 mytilus gal
3	146.5	43.2	73	5 Q9BLJ3	Q9BLJ3 ornithodor
4	140	41.3	73	5 Q9BLJ4	Q9BLJ4 ornithodor
5	108	31.9	57	5 Q9Y0B0	Q9Y0B0 mytilus gal
6	90	26.5	91	5 Q9N6G1	Q9N6G1 mesobuthus
7	82.5	24.3	85	5 Q9NJC7	Q9NJC7 mesobuthus
8	79.5	23.5	79	5 Q9G0A9	Q9G0A9 oryctes rhi
9	77	22.7	57	5 Q9GYU6	Q9GYU6 aedes allop
10	75.5	22.3	85	5 Q9NJP8	Q9NJP8 mesobuthus
11	75.5	22.3	85	5 Q9GYX2	Q9GYX2 mesobuthus
12	71.5	21.1	85	5 Q9BLM4	Q9BLM4 androctonus
13	70	20.6	102	5 Q61721	Q61721 anopheles q
14	70	20.6	103	5 Q9Y0B1	Q9Y0B1 mytilus gal
15	69.5	20.5	72	10 Q9FR81	Q9FR81 pisum sativ
16	68.5	20.2	85	5 Q9GNG8	Q9GNG8 mesobuthus

17	67.5	19.9	79	10 Q39807	Q39807 glycine max
18	66.5	19.6	85	5 Q9GUA7	Q9GUA7 mesobuthus
19	66.5	19.6	85	5 Q9SP69	Q9SP69 mesobuthus
20	65.5	19.3	195	11 Q91VZ7	Q91VZ7 mus musculu
21	64	18.9	96	5 Q77217	Q77217 aedes allop
22	64	18.9	98	5 P91793	P91793 aedes aegyp
23	64	18.9	98	5 Q963E9	Q963E9 aedes aegyp
24	63.5	18.7	81	10 Q9XGD9	Q9XGD9 zea mays (m
25	63.5	18.7	83	5 Q9BK52	Q9BK52 acalolepta
26	63	18.6	73	5 Q9FQ14	Q9FQ14 citrus para
27	62	18.3	87	10 Q948T4	Q948T4 pyrus pyrif
28	62	18.3	490	16 Q9PKG3	Q9PKG3 chlamydia m
29	62	18.3	990	13 Q91803	Q91803 xenopus lae
30	62	18.3	4599	11 Q9J118	Q9J118 mus musculu
31	61.5	18.1	85	5 Q9Y0B8	Q9Y0B8 mesobuthus
32	61.5	18.1	154	12 Q84631	Q84631 paramescium
33	61	18.0	39	5 P82380	P82380 stomoxys ca
34	61	18.0	643	11 Q9ET61	Q9ET61 rattus norv
35	61	18.0	643	11 Q9J1Z6	Q9J1Z6 rattus norv
36	61	18.0	644	11 Q89103	Q89103 mus musculu
37	60.5	17.8	81	5 Q94435	Q94435 centrurorde
38	60.5	17.8	85	5 Q9YX87	Q9YX87 mesobuthus
39	60.5	17.8	85	5 Q95WX6	Q95WX6 mesobuthus
40	60.5	17.8	126	5 Q45448	Q45448 caenorhabdi
41	60.5	17.8	652	4 Q9NYP3	Q9NYP3 homo sapien
42	60.5	17.8	652	4 Q00274	Q00274 homo sapien
43	60	17.7	85	5 Q9Y1U3	Q9Y1U3 mesobuthus
44	60	17.7	85	5 Q9UAC8	Q9UAC8 mesobuthus
45	60	17.7	142	4 Q9NZK7	Q9NZK7 homo sapien

ALIGNMENTS

RESULT 1  
ID Q9Y0A9 PRELIMINARY; PRT; 81 AA.  
AC Q9Y0A9;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE DEFENSIN MGD-2 PRECURSOR.  
OS Mytilus galloprovincialis (Mediterranean mussel).  
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;  
OC Mytiloidea; Mytilidae; Mytilus.  
OX NCBI\_TaxID=29158;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HEMOCYTE;  
RX MEDLINE=20033584; PubMed=10564642;  
RA Mitta G., Vandenbulcke F., Hubert F., Roch P.;  
RT "Mussel defensins are synthesised and processed in granulocytes then released into the plasma after bacterial challenge.";  
RL J. Cell Sci. 112:4233-4242(1999).  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- MISCELLANEOUS: THE PI OF THIS PROTEIN IS 8.92.  
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
DR EMBL: AF162338; AAD45118.1;  
DR InterPro: IPR001542; Arthro\_defensin.  
DR InterPro: IPR003614; Knot1.  
DR Pfam: PF01097; Arthro\_defensin; 1.  
DR SMART: SM00505; Knot1; 1.  
DR PROSITE: PS00425; ARTHROPOD\_DEFENSIN; UNKNOWN\_1.  
KW Defensin; Antibiotic; Fungicide; Amidation; Signal.  
FT SIGNAL 1 21 POTENTIAL.  
FT PROPEP 22 60 POTENTIAL.  
FT CHAIN 61 81 DEFENSIN MGD-2.  
FT MOD\_RES 60 60 AMIDATION (POTENTIAL).  
FT DISULFID 25 46 BY SIMILARITY.  
FT DISULFID 31 54 BY SIMILARITY.  
FT DISULFID 35 56 BY SIMILARITY.  
FT DISULFID 42 59 BY SIMILARITY.  
SQ SEQUENCE 81 AA; 9087 MW; 0EEF7AB5CF6DCE53 CRC64;

Query Match	45.6%	Score 154.5;	DB 5;	Length 81;
Best Local Similarity	51.7%	Pred. NO. 1.3e-12;		
Matches 30; Conservative		7; Mismatches 18;	Indels 3;	Gaps 3;

**Qy** 5 AIIIVLVAFCILEDGIVEAGFGCPFNACKCHRRCKSIRRR-GGCRGTFRITCVYR 61  
| : :: | : : | : : | : : | : : | : : | : : | : : | : : | : :  
**Dd** 3 AAFLVLLVGLCIMTD-VATAGFGCPNNA-CHOHCKSIRGYCGGYCAGWFLRCTCYR 58

RESULT	2	
Q9U600		
ID	Q9U600	PRELIMINARY; PRT; 82 AA.
AC	Q9U600;	
DT	01-MAY-2000 (TREMBLrel. 13, Created)	
DT	01-MAY-2000 (TREMBLrel. 13; Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19; Last annotation update)	
DE	ANTIMICROBIAL PEPTIDE MGD2B.	
OS	Mytilus galloprovincialis (Mediterranean mussel).	
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;	
OC	Mytiloidea; Mytilidae; Mytilus.	
OX	NCBI_TaxID=29158;	

[1]  
SEQUENCE FROM N.A.  
Mitta G., Hubert F., Dvirya E.A., Boudry P., Roch P.;  
"Mytilin B and MGD2: Two antimicrobial peptides of marine mussels:  
gene structure and expression analysis.",  
Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
EMBL: AF17539; AA052660.1; --  
InterPro: IPR001542; Arthro\_defensin.  
InterPro: IPR001545; Glyco\_hormone\_beta.  
Pfam: PF01097; Arthro\_defensin; 1.  
PROSITE: PS00425; ARTHROPOD\_DEFENSINS; UNKNOWN\_1.  
PROSITE: PS00261; GLYCO\_HORMONE\_BETA\_1; UNKNOWN\_1.  
SEQUENCE 82 AA; 9307 MW; 608F0039PF675ECE CRC64;

Query Match 44.8%; Score 152; DB 5; Length 82;  
Best Local Similarity 48.3%; Pred. NO. 2.9e-12;  
Matches 28; Conservative 8; Mismatches 20; Indels 2; Gaps 2;

**Qy**      5 AIIIFVLVAFCLIEDGIVEAGFCPENAGKCHRHCKSIARR-GGFCRGTFRTTCVCYR 61  
       I- : :: I :: : I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I :: I ::  
**Db**      3 AFVLLVVGLCINMDVATAGFGCPNYYA-CHOHCKSIRGYCGYG/CASWFLRLRCTCYR 59

RESULT	3			
Q9BLJ3				
ID	Q9BLJ3	PRELIMINARY;	PRT;	73 AA.
AC	Q9BLJ3;			
DT	01-JUN-2001	(TREMBlrel. 17, Created)		
DT	01-JUN-2001	(TREMBlrel. 17, last sequence update)		
DT	01-OCT-2001	(TREMBlrel. 18, last annotation update)		
DE	DEFENSIN A.			
DE	OMDER-A.			
GN	Ornithodoros moubata (Soft tick).			
OS	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
OC	Parasitiformes; Ixodida; Argasidae; Ornithodoros.			
OX	NCBI_TaxID=6938;			

Query Match 43.2%; Score 146.5; DB 5; Length 73;

		Best Local Similarity	57.5%	Pred. No. 1.3e-11;
		Matches	23; Conservative	6; Mismatches 10; Indels 1; Gaps 1;
QY	22	VEAGFGCFNAGKCHRHCKSIR-RGGFGCRGTFRTTCVCY	60	
		: : : : :   : : : : :   : : : : :   : : : : :		
Db	34	VRRYGCFPNQYCHSHCSGIRGKYGKCGKCFKOTCKCY	73	

RESULT 4  
Q9BLJ4  
ID Q9BLJ4

```

ID          Q9BLJ4      PRELIMINARY;      PRT;      73 AA.
AC          Q9BLJ4;
DC          01-JUN-2001 (TrEMBLrel. 17, Created)
DT          01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT          01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE          DEFENSIN B.
DE          OMDEF-B.
GN          Ornithodoros moubata (Soft tick).
OC          Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC          Parasitiformes; Ixodida; Argasidae; Ornithodoros.
OX          NCBI_taxid=6938;
RN          [1]
RP          SEQUENCE FROM N.A.
RA          Nakajima Y.;
RA          "CDNA CLONING OF TICK DEFENSIN."
RT          Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RL          EMBL: AB041815; BAB14027.1; -.
DR          InterPro: IPR001542; Arthrodefensin.
DR          Pfam: PFO1097; Arthrodefensin; 1.
DR          PROSITE: PS00425; ARTHROPOD_DEFENSIN; UNKNOWN_1.
DR          SEQUENCE 73 AA; 8290 MW; DED8E90B06C9448D CRC64;
SQ

```

Query Match 41.3%; Score 140; DB 5; Length 73;  
Best Local Similarity 40.3%; Pred. No. 9e-11;  
Matches 27; Conservative 11; Mismatches 19; Indels

Qy	4	I A I F I V L V A F C I L E - D G I - - - - - V E A G F G P F N A G K C H R H C K S I R - R R G G F C R T F F	53
Dd	7	V A L V A L A V A T M A Q E V H D D V E E Q S V P R V R R Y G G P F N Q Y O C H S H C R G I R G Y K G G Y C T G R F	66
Qy	54	R T T C V C Y 60	
Dd	67	K O T C K C Y 73	

RESULT	5	
Q9Y0B0		
ID	Q9Y0B0	PRELIMINARY;
AC	Q9Y0B0;	PRT;
DT	01-NOV-1999	(TrEMBLrel. 12, Created)
DT	01-NOV-1999	(TrEMBLrel. 12, Last sequence update)
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)
DE	MGDI ANTIMICROBIAL PEPTIDE (FRAGMENT).	
OS	Mytilus galloprovincialis (Mediterranean mussel).	
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;	
OC	Mytilodea; Mytilidae; Mytilus.	
OX	NCBI_TaxID=29158;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Mitta G., Hubert F., Roch P.;	
RT	"Defensins: involvement in mussel defense.";	
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; AF162337; AAD45117.1;	
DR	InterPro; IPR001542; Arthro_defensin.	
DR	Pfam; PF01097; Arthro_defensin; 1.	
DR	PROSITE; PS00425; ARTHROPOD_DEFENSINS; 1.	
DR	NON_TER	1
FT	SEQUENCE	57 AA; 6658 MW; 7B86B8623C979326 CRC64;

Query Match 31.9%; Score 108; DB 5; Length 57;  
Best Local Similarity 60.0%; Pred. No. 9.5e-07;



DR	InterPro; IPR002061; Scorpion_toxin.
DR	InterPro; IPR001219; Toxin.
DR	Pfam; PF00537; toxin_3; 1.
DR	PRINTS; PR00284; TOXIN.
DR	ProDom; PD000908; Scorpion_toxin; 1.
DR	SMART; SM00505; Knot1; 1.
SQ	SEQUENCE 85 AA; 9588 MW; 416CB3D72A8927ED CRC64;
Query Match 22.3%; Score 75.5; DB 5; Length 85;	
Best Local Similarity 26.6%; Pred. No. 0.022;	
Matches 17; Conservative 13; Mismatches 27; Indels 7; Gaps 3;	
QY	3 STAIIFVLVAFCILEDGIVEAGFCGFENACK---CHRHCKSIIRRGGR--GTFRFTTC 57 !:::   :    :   :                         :
Db	8 SLALLMTGVE--SVKDGVIADNRNCFPGCRNAYCDGECKKNRAESGYCQWASKYGNC 65 !:::   :    :   :                         :
QY	58 VCVR 61
Db	II:
Db	66 WCYK 69
RESULT 11	
OQGX2	PRELIMINARY; PRT; 85 AA.
ID OQGX2 AC	
DT 01-WAR-2001 (Tremblrel. 16, Created)	
DT 01-WAR-2001 (Tremblrel. 16, Last sequence update)	
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)	
DE ALPHA_TOXIN 1 PRECURSOR.	
OS Mesobuthus martensii (Manchurian scorpion) (Buthus martensii).	
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;	
OC Buthoidea; Buthidae; Buthus.	
OX NCBI_TaxID=34649;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Ye J., Chen J., Zuo X., Ji Y.;	
RT "Cloning and characterization of cDNA sequences encoding two novel	
RT alpha toxin precursors from the Chinese scorpion Buthus martensi	
RT Karsch.";	
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.	
DR EMBL: AF288607; AAG00580.1; -	
DR HSP: p17728; l1QI	
DR InterPro; IPR003614; Knot1.	
DR InterPro; IPR002061; Scorpion_toxin.	
DR InterPro; IPR001219; Toxin.	
DR PRINTS; PR00284; TOXIN.	
DR ProDom; PD000908; Scorpion_toxin; 1.	
DR SMART; SM00505; Knot1; 1.	
KW Signal.	
FT SIGNAL.	1 19 POTENTIAL.
FT CHAIN	20 85 ALPHA_TOXIN 1.
SQ SEQUENCE 85 AA; 9366 MW; 94ED463D6193FFEF CRC64;	
Query Match 22.3%; Score 75.5; DB 5; Length 85;	
Best Local Similarity 26.6%; Pred. No. 0.022;	
Matches 17; Conservative 12; Mismatches 28; Indels 7; Gaps 3;	
QY	3 STAIIFVLVAFCILEDGIVEAGFCGFENACK---CHRHCKSIIRRGGR--GTFRFTTC 57 !:::   :    :   :                         :
Db	8 SLALLMTGVG--SVRDGIADNRNCFPGCRNAYCDGECKKNRAESGYCQWASKYGNC 65 !:::   :    :   :                         :
QY	58 VCVR 61
Db	II:
Db	66 WCYK 69
RESULT 12	
OQBML4	PRELIMINARY; PRT; 85 AA.
ID OQBML4 AC	
DT 01-JUN-2001 (Tremblrel. 17, Created)	



```

Query Match      21.1% Score 71.5 DB 5 Length 85;
Best Local Similarity 27.3% Pred. No. 0.072;
Matches 18; Conservative 12; Mismatches 29; Indels 7; Gaps 3;

QY 1 KMSIAITFLVAFCLDGVAGCGPPNAGK---CHRHCKSIIRRGGFCR--GTFTT 55
      |||::||| | | | | | | | | | | :
Db 6 MISLALFETGVE--SKKDGYIVDDKNCTFFCGRNAYCNDECKKGAESGYCOWASPYGN 63
      ||| | | | | | | | | | | | | |

QY 56 TCVCYR 61
      | | |
Db 64 ACYCYN 69

```

Query Match	20.6%	Score 70;	DB 5;	Length 102;
Best Local Similarity	41.7%	Pred. No. 0.13;		

```

Query Match      20.68; Score 70; DB 5; Length 103;
Best Local Similarity 33.9%; Pred. No. 0.14;
Matches 20; Conservative 8; Mismatches 23; Indels 8; Gaps 3;

Qy 1 MKSAIIIFIVLAFCEILDGIVEAGFGPPFNAGKCRHHRCKSIIRRGGFCRGTFRTTCVC 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MKAAVIIAIIAIVAIL----AVHAEAC-- --ASRCKGHCRRA--RRCGYVSVLYRGRCVC 51

```

```

RESULT 15
Q9FR81 PRELIMINARY; PRT; 72 AA.
AC Q9FR81:
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DISEASE RESISTANCE RESPONSE PROTEIN 230.
GN DR230.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. GREENFEAST;
RA Savenstrand H., Brosche M., Strid A.;
RT "Stress-induced disease resistance response protein 230 cDNA from
RT Pisum sativum cv. Greenfeast";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF139018; AAG43285.1; -.
DR InterPro: IPR002118; Gamma-thionin.
DR InterPro: IPR003614; Knott1.
DR Pfam: PF00304; Gamma-thionin; 1.
DR ProDom: PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knott1; 1.
SQ SEQUENCE 72 AA: 8000 MW: 00F9B60E5FF0D920 CRC64;

```

	Query Match	20.5%;	Score 69.5;	DB 10;	Length 72;
	Best Local Similarity	29.9%;	Pred. No. 0.11;		
	Matches	20;	Conservative 10;	Mismatches 24;	Indels 13; Gaps 3;
QY	2 KSIATIFVLVAFCLIEDGIVEAGCGCFNNAGK-----CHRHCKSIRRRG---GFGR 50     : : :   :          :				
Db	4 KSLACLFLLLVLFAIQEIVVSEANTCENTLAGSYKVCFGGCDRHCT--QEGAISGRCR 61				
QY	51 GTFRTTC 57 				
Db	62 DDFRCWC 68				

Search completed: September 17, 2002, 15:45:33  
Job time: 129 sec